Journal of Engineering Science and Technology
Vol. 16, No. 6 (2021) 4766 - 4778

INFLUENCE OF GROWTH PARAMETERS ON THE CROP YIELD PERFORMANCE OF HYDROPONIC SPINACH (SPINACIA OLERACEA L.) USING CORRELATION AND REGRESSION MODELS

SRIVANI P.[1,*], YAMUNA DEVI C. R.[2], MANJULA S. H.[3]

[1]BMS Institute of Technology and Management, Bengaluru, India
[2]Dr. Ambedkar Institute of Technology, Bengaluru, India
[3]University of Visvesvaraya College of Engineering, Bengaluru, India
*Corresponding Author: srivanicse@bmsit.in

Abstract

Different growth parameters highly influence plant growth dynamics of the hydroponic system. Several predictive models can be applied to understand the plant growth dynamics in a hydroponic system. The spinach crop's plant growth cabinet had been designed to measure pH, EC, water temperature, air temperature, and humidity in a floating raft system. The target variables, the number of leaves, and length of a plate (in cm), measured day-wise, were considered in predicting the crop yield. Three correlation models were considered to analyse the association between multivariate input parameters with target variables. The results show that the Electrical Conductivity (EC) of the nutrient solution has a strong positive correlation. And to determine the crop yield, several Machine leaning based regressor models (Random Forest, Gradient Boost, K-Nearest Neighbors, Bayesian Ridge) were implemented. The models' performance was evaluated and compared based on accuracy score and RMSE. The results showed better prediction with the Random Forest regressor model with accuracy greater than 98%. Hence, with this study, the yield of indoor hydroponic plants could be improved by controlling pH and EC which has significant effect.

Keywords: Correlation analysis, Electrical conductivity, Hydroponic, Plant growth, Prediction, Regression model.

1. Introduction

Hydroponics is a soilless agricultural system where plant growth depends on the nutrients added. There is a lot of demand for locally grown chemical-free, organic, safe produce crop production in recent years. Through smart hydroponic cultivation, this can be achieved with better efficiency. Recent advancements on the Internet of Things (IoT), Information and Communication Technology (ICT), Machine Learning, and Artificial Intelligence (AI) have a great potential to address a few of the environmental, technical, and economic challenges in the agricultural sector. Smart farming with a controlled environment hydroponic system is emerging as one solution towards enhanced farm productivity and efficiency. As per the statistics reports on the global population, there is more than 25% higher yield with a controlled environment hydroponic system than the traditional growing system [1]. This type of system has minimal impact on climatic conditions, and different varieties of crop production all year round can be set up. The comparative study [2] with greenhouse and indoor hydroponic has illustrated a decreased plant growth rate and might also need a frequent refill of water for indoor cultivation.

Spinach (*Spinacia oleracea L.*) is a nutrient-dense food and is sensitive for growth in a cool climate. This crop is a prime source of antioxidants, Vitamin A, C, E, K, and also rich in iron, calcium, and magnesium. To enhance the quality and productivity of the crops, an alternate source for freshwater explored in the experiment for the spinach plant [3] were different seawater treatments, which had a positive impact on growth and improved sodium content. However, this is more suitable in coastal areas where seawater is available.

The relative growth rate of spinach was greatly affected by nutrient solution temperature with the correlation of 0.61, compared to other environmental parameters by implementing a multiple linear regression model [4]. In this study, with a lower pH (<4), there was a negative influence on plant growth, and the environmental temperature had an influenced metabolism of sugar in the plant. Also, the plant's nutrient uptake is directly proportional to the pH of the solution. The plant growth process under artificial lightning [5] will absorb more nitrate content for leafy vegetables. The physiological factors on plant growth had a positive effect, and plants grown in high EC and low pH nutrient solution had higher chlorophyll content in leaves and have faster growth rate in an ebb-flow system [6]. The hypothesis of this study was to identify the parameters which have null and alternate hypothesis states with plant growth rate using correlation models. This study aims to predict the yield with greater accuracy than an existing system with an efficient and best regressor model with a lower error rate (<2).

Related work

Crop yield predictions with mathematical non-destructive models and statistical analysis are more efficient and convenient ways. Solis-Toapanta et al. [2] have experimented on basil and compared the growth rate concerning greenhouse and indoor environments. The decline in nutrient uptake and growth rate was depicted indoors. The addition of fertilizers in hydroponic may create fluctuations in the EC level. The state of the art of literature related to existing systems and the models implemented on hydroponic plant growth is shown in Table 1.

Table 1. Few of the existing prediction models implemented on different crops.

Ref.	Hydroponic plant	Input parameters	Target parameters	Model	Accuracy
[7]	Cucumber	Leaf length and width	dry weight and fresh weight	Regression Models	97%
[8]	Lettuce	solar irradiance, temperature and vapour-pressure deficit	Dry-weight and leaf area index	Dynamic Bayesian Network	96%
[9]	Lettuce	Dissolved-oxygen, nutrient solution temperature, EC, pH, temperature & humidity	Photosynthetic rate	Artificial Neural network	91%
[10]	Tomato	Vapour pressure deficit, CO_2, temperature	Yield(in Kg)	Evolving Fuzzy neural Network	90%
[11]	Lettuce	Leaf images	Leaf movement	multi-plant imaging model using Deep flow	-
[12]	Mustard	Plant images	Age and fresh weight	Artificial neural network	95%-96%

This study has involved automatic monitoring and collection of environment parameters and nutrient solution parameters using IoT technology. The present study analyses the correlation between the hydroponic parameters (pH, EC, and nutrient solution temperature) & environmental parameters (room temperature and humidity) with plant growth parameters such as height and number of leaves using correlation models. Further, to predict the plant growth, the five regression models are proposed and compared to choose the best model based on the accuracy score. All these models are implemented on a multivariate dataset considering five input variables and one target variable. Several regression models are implemented to know the best performance model with a multivariate dataset.

2.Material and Methods

2.1. Plant growing system

The plant growing system chosen for growing hydroponic spinach plant is DWC which is shown in Fig. 1.This research work was experimented on spinach and was developed over 75 days from 21st July to 4th October 2020. The plants were transplanted to the deep floating raft system or DWC system after two weeks of seedling in coco-peat. The recommended pH of nutrient solution for spinach is 5.5-6.5 [13]. The plant growth was monitored automatically using sensors and is uploaded to the cloud environment.

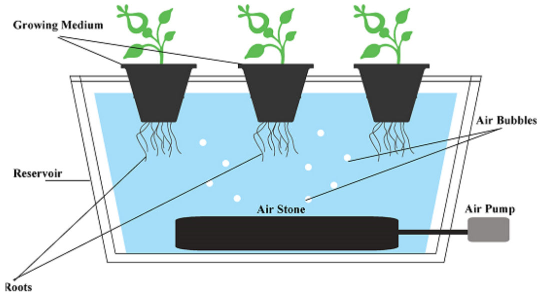


Fig. 1. Structure of deep-water culture system.

With the integration of IoT technology, pH sensor, EC sensor, environment monitoring sensors (DHT22), and nutrient solution temperature sensor (DSB18B20) are used to sense the data and upload in Google Sheets Day wise. The parameters automatically measured as inputs from the sensors were nutrient temperature (0C), Electrical conductivity (ppm), pH of nutrients (0-14 range), and room temperature (0C), and humidity (%). The target variables identified were the number of leaves and the length/height of leaves (in cm). The observed data were analysed to understand the association between parameters and target variables. Several machine learning regression models were implemented to predict the height and number of leaves of spinach plants. The plant monitoring hydroponic system was cultivated under Grow LED Lights with 12/12 (light/dark) hours for the photosynthesis process with a light intensity of 27 watts and 8000 kelvins.

Figure 2 shows the plant growth of spinach in an indoor environment. Figure 2(a) demonstrates the plant growth in a seedling tray during the germination stage. Figure 2(b) shows the growth of plant of day five after transplantation into the hydroponic raft system.

(a) Germination under coco-peat. **(b) Day 5 after transplantation.**

Fig. 2. Hydroponic plant growth of spinach in indoor environment.

2.2. Methodology

Prediction/forecasting crop production is one of the hot throbs of research trending for maximizing crop production and minimizing resources. With technology, an automated intelligent system can be built for socio-economic concerns in the precision agricultural system. Studying the relationship between the parameters is helpful in understanding the dependency before we attempt to predict the plant growth dynamics. Multivariate regression analysis models were implemented to find the significance of correlation and prediction. The correlation of hydroponic parameters with leaf area and biomass is implemented using the Verhulst model [14], and a linear regression model would be more suitable to analyse the linear association between plant growth dynamics and pH and EC [15]. In this research, the correlation models are proposed to analyse the linear and non-linear association with plant growth dynamics. Further, to predict the crop yield on the spinach growth dataset, the regressor models were proposed and implemented. The models were developed using Python language on a Jupyter Notebook platform which includes several libraries like Tensorflow, sci-kit, matplolib, pandas, etc.

2.2.1. Correlation analysis models

Correlation analysis is a statistical method that shows how strongly or weakly the variables are interdependent. Analysing the statistical correlation between the parameters that affect the dynamics of plant growth is very important, which ranges between -1 to 1. It estimates how a change in one variable is related to a change in another variable base on hypothesis testing. The correlation test with null hypothesis determines no association between the measured parameters. Similarly, a correlation test with the alternative hypothesis defines that there exists a relationship between the measured parameters.

Pearson correlation analysis determines the linear relationship between the parameters, whereas Spearman and Kendall's Rank correlation identifies both linear and non-linear monotonic relationships based on rank. The correlation analysis is implemented with multiple parameters with these three models where the input parameters (X) are hydroponic parameters & environmental parameters. The target variables (Y) are calculated individually with length and height of leaves. The association is between each X parameter and each Y parameter to assess weak and strong correlation.

The Pearson correlation coefficient between two variables is calculated using the Eqs. (1) and (2). The variance of X or Y determines how the data is distributed across the mean value.

$$R_{\text{pearson}}=\frac{\text{covariance}(X,Y)}{\sqrt{(var(X))}*\sqrt{(var(Y))}} \quad (1)$$

$$\text{covariance}(X,Y)=\sum_{j=1}^{N}\ (X_j-\overline{X})*(Y_j-\overline{Y}) \quad (2)$$

Spearman rank correlation method assesses how good an arbitrary monotonic association would describe the relationship between the parameters [16]. Higher the absolute value stronger the correlation and ranges between -1 to +1. The Spearman calculates based on the rank using Eq. (3), where $rg(X_i)$, and $rg(Y_i)$ represents converted rank scores for the given n observations. The correlation is measured when the observations do not have tied ranks.

$$\rho_{\text{spearman}}=1-\frac{6\sum(rg(X_i)\text{-}rg(Y_i))^2}{N*(N^2\text{-}1)} \quad (3)$$

Kendall correlation analysis is more efficient and robust than Spearman which can be used for smaller samples [16], which is non-parametric and has better statistical characteristics. Positive values propose that the higher the total worth more grounded the relationship between the factors. Negative values recommend that higher estimations of one are related to the lower estimates of the other. To calculate the rank and associativity, the Eq. (4) is referred to as τ_{bken}.

$$\tau_{\text{bken}}=\frac{\text{number of concordant pairs} - \text{number of discordant pairs}}{\text{number of concordant pairs} + \text{number of discordant pairs}} \quad (4)$$

2.2.2. Regressor models for prediction

Multiple regression models learn about the relationship between the various independent predictor parameters and a single dependent variable. Apart from correlation analysis, the multiple regression models aid in predicting the output parameter. Multivariate regression models are compared for predicting plant growth and height and the influence of several factors on the plant. The efficient crop growth is analysed using the mean squared error and best prediction model with most minor errors.

Random Forest Regressor (RFR) model is normally characterized by developing classifying decision trees depending upon the various sub-samples of featured data. The hyperparameter of the model is defined with 500 *n*_estimators (number of decision trees) by choosing random of n samples that perform prediction by aggregating. For a given randomly chosen input vector $f(\theta)$ in terms of X, the random forest or predictor tree $h(X, \theta)$, the samples are independently selected from the joint distribution of (X, Y) [17]. The model performance is measured using mean square error using the following Eq. (5).

$$E_{X,Y}(Y - \bar{h}(X)^2) = E_{X,Y}(Y - E_\theta * h(X,\theta))^2 \tag{5}$$

where the aggregation of all the decision trees is used to predict the values.

Gradient Boosting Regressor (GBR) model is an ensemble supervised learning. It depends on the intuition that it fits with the best possible additional model when merged with past models. Thus by minimizing the prediction error or specific loss function over the training dataset [18]. This will take the form of a gradient forward stage model that sets the target outcomes by applying a fixed regression tree technique. The model is set to define 500 *n*_estimators. As the number of stages in GBR is set to 500, the weights are minimized after computing the loss function. Initially, the weights for all the samples have identical weights; later, to predict the new observations, the model considers the weighted aggregation of predictions of earlier tree models.

The *K*-nearest Neighbors Regressor (KNR) model works based on the number of *K*-nearest neighbours to limit false predictions that estimate continuous variables. A range of different *k*-values is considered to choose an optimal value. The number of neighbours selected to implement on this dataset is five (k=5), which predicts the new value based on these neighbours. For a given set of X and Y values from the dataset, a Euclidean distance is used to calculate between the data points shown in Eq. (6).

$$D(X,Y) = \sum_{j=1}^{N} (X_j - Y_j)^2 \tag{6}$$

Finally, for all *j*-histories, the distances are ranked, and the lowest feature distance vector is averaged to predict the parameter. The model is set to neighbours of 5 for the plant growth analysis. The performance of the algorithm depends on how well we select the trained data.

The Bayesian Ridge Regressor (BRR) model permits natural mechanisms to endure inadequate information or ineffectively disseminated information by formulating linear regression with probability distributors instead of value estimates [19]. The Bayes theorem is used to predict subsequent distribution. Mathematically, ridge regression is represented based on L_2 constrained least squares, using Eq. (7), where λ is considered as a random variable for prediction, which is a regularization term, for a given vector of β.

$$\hat{\beta} = arg_\beta min(||Y - \mu 1 - X\beta||^2 + \lambda||\beta||^2) \tag{7}$$

The ridge multi-regression model has high co-linearity among the target/predictor variables. The bigger the estimation of λ, the more the parts of β are contracted towards zero.

2.2.3. Growth co-relation and prediction model using correlation and regressor models

The growth correlation and prediction models are implemented using correlation and few regression models. The algorithms are applied to the trained dataset and later tested with the test dataset. Table 2 demonstrates the pseudocode for the correlation and prediction model for plant growth. To analyse the best outcome, these models are compared with accuracy scores and standard errors.

Table 2. Pseudo code for growth correlation and prediction model for multivariate dataset.

```
1.Choose the dataset for analysing the correlation between multivariate parameters for
  analysing plant height(in cm) and number of leaves
2.Inputs are Humidity, Air Temperature, Nutrient Solution temperature, pH and EC from
        a given dataset(data) for n=75
3.While(true)
    Input method for correlation analysis
    Call function correlation( data, method)
    If (method== "none") then
      Break
    end if
  end while
4. Split the dataset features into training and testing set of a ratio (8:2)
5. Call function regression(training set, testing set)
6. Find the best accuracy score among the different models
7. Best_accuracy_score= Max(score of all the models)
8. Goto step 2 and repeat the steps for analysing the growth w.r.t number of leaves
9. End

Function 1
Correlation( data, method)
    If (method = ="pearson") then
      Compute the correlation matrix w.r.t input parameters and Height of plant using
      Pearson correlation co-efficient
    else if (method= ="spearman") then
      Compute the correlation matrix w.r.t input parameters and Height of plant using
      Spearman  correlation
    else
      Compute the correlation matrix for Kendall method w.r.t input parameters and
Height of plant
    end if
end function

Function 2
Regression(training set, testing set)
    For each type of regressor model
          Define and set the hyperparameters of each model
          Compute the accuracy score and root mean square error
          Return accuracy score , root mean square error
 end function
```

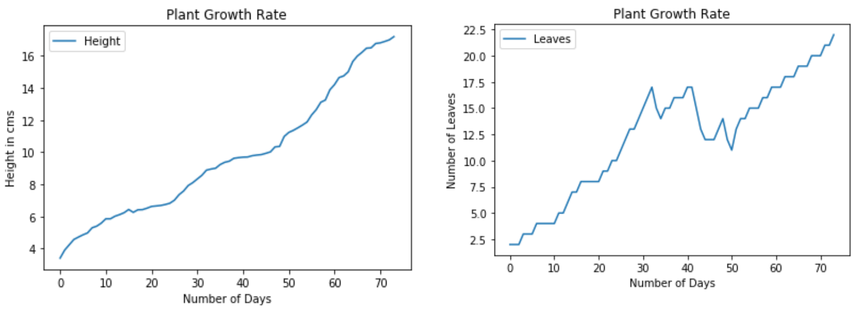
3.Results and Discussions

The experiment results with the DWC system on spinach growth rate significantly impact the pH parameter. Figure 3(a) shows an effective growth when pH was maintained or controlled by adding basic/acidic solution when it was out of range (5.5-6.5), and EC was maintained around 1000-1500 range. Figure 3(b) picturizes the decline in plant growth when pH was not controlled and was not within the optimal range, due to which the absorption of nutrients was insufficient. The sign of deficiency of nutrients can be observed in Fig. 3(b), which shows pale yellow leaves and the falling of leaves. This study highlights a more substantial influence of pH and electrical conductivity of the nutrient solution.

(a) Effective plant growth. **(b) Decline in plant growth.**

Fig. 3. Plant growth dynamics with controlled and uncontrolled pH.

The plant growth rate is analysed using the height of the plant in cms and the number of leaves. In general, several researchers have predicted the yield in terms of fresh weight, dry weight, leaf width, length, and stem diameter. The parameters have greatly affected the height and number of leaves which is visualized in Fig. 4. Figure 4(a) shows growth with height over a period of time. As observed in Fig. 4(b), there was a decline in plant growth during the days between 35 to 50 due to the uncontrolled pH value(above 8) and the height was almost stable shown in Fig. 4(a).



(a) Growth in terms of height. **(b)Growth in terms of number of leaves.**

Fig. 4. Plant growth rate of spinach with time.

Model performance metrics

In this study, analysing the correlation between the parameters and the prediction of crop yield was done using correlation and regression models. Since the crop yield harvesting period is significantly less for leafy vegetables compared than other veggies, the number of days considered is 72. Hence, the size of the dataset, which is a day-wise accumulation to analyse the plant growth from the stage of transplantation till harvesting, is 72 rows. The symptoms of nitrogen deficiency were suspected in this study as the leaves started turning pale yellow. The observed characteristic of this study demonstrated a stunted growth, and the pH of the nutrient solution was increasing above the optimal value. The correlation coefficient is compared with different correlation methods for linear and non-linear dependencies on multivariate parameters.

Figures 5 and 6 demonstrate the correlation co-efficient with the crop growth parameters using Pearson, Spearman, and Kendall methods for crop growth analysis. This summary of statistics of the features is analysed concerning to plant height, as shown in Fig. 5. As per the observations, there were few instances affected the crop yield in terms of height and number of leaves. The humidity on the plant growth rate for height did not affect (Null Hypothesis) and the other parameters determine alternate hypothesis with either positive or negative correlation. The Electrical conductivity of the nutrient solution has the strongest positive correlation, which determines that as EC increases, the plant growth rate also increases with time. The plant growth is linearly related to decreasing relationship with pH value based on the Pearson correlation model. For Spearman and Kendall models, the positive correlation with environmental parameters, EC, and nutrient solution temperature, indicates that the rank increases with plant growth. The above-mentioned non-parametric models (Spearman and Kendall) evaluates the association based on ranks, where the association is not dependent on assumptions of distribution.

Similarly, the crop growth measured in terms of the leaves' count is analysed using all three correlation coefficient models. EC of the nutrient solution has a positive correlation concerning all three models, which directly impacts plant growth rate. The parameters are monotonically and linearly related to plant growth. The same is shown in Fig. 6 to demonstrate the correlation analysis with leaves.

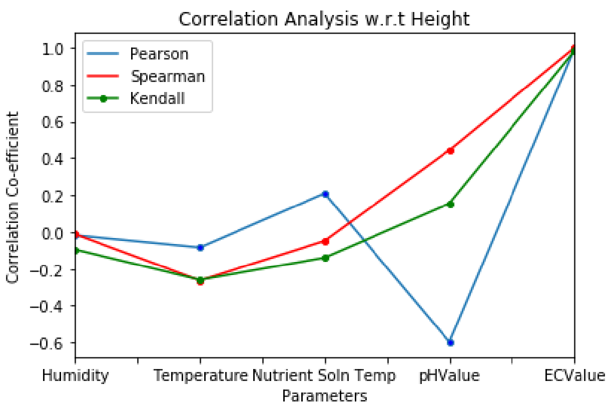


Fig. 5. Correlation analysis- height of the plant (in cm).

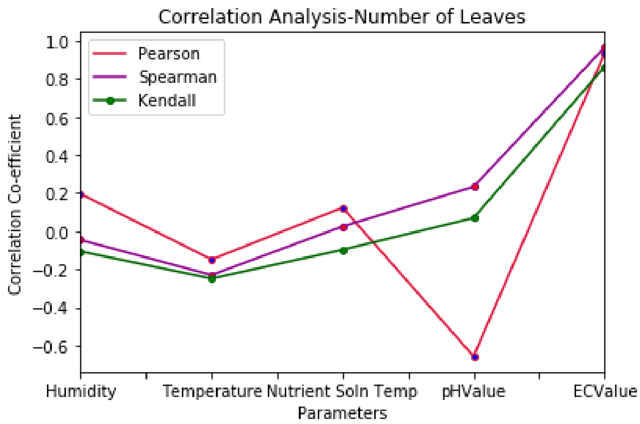


Fig. 6. Correlation analysis- number of leaves of the plant.

The results show that concerning the Pearson Co-relation coefficient, a negative correlation occurs with the nutrient solution's parameter pH that significantly affects the plant growth. This correlation depicts that when a pH was high(>7), there was a decline in plant growth and the leaves. When pH was controlled, then the growth was average and good. As discussed in the literature, the optimal pH value range on leafy vegetables is 5.5-6.5. With lower pH(<4), suppression of plant growth and nutrient disorder exists [13]. And with a higher pH value, the present study evidence that the growth rate in terms of height and number of leaves was reduced.

The next stage is to do prediction using regression models. The regression models like Random Forest Regressor, Gradient Boosting Regressor, *K*- Nearest Neighbors Regressor, and Bayesian Ridge Regressor was implemented to predict the plant height and number of leaves. The prediction performance was evaluated using Root Squared error (R^2) or accuracy score. The standard error of the regression is calculated using the root mean square error (RMSE) measured for the height and leaves parameter. Figure 7 shows the predicting capability of all regression models for the testing data for crop growth dynamics for measuring height and leaves. The comparison of accuracy scores and RMSE of the models calculated from the prediction for the plant's height and a number of leaves are shown in Table 3.

Table 3. Summary of regressor models' performance.

Model	Accuracy Score		RMSE	
	Plant Height	Number of Leaves	Plant Height	Number of Leaves
RFR	98.3	92.9	0.5	1.3
GBR	97.0	91.9	0.6	1.4
KNR	95.6	92.8	0.7	1.3
BRR	85.8	65.6	1.3	2.9

After the analysis, the best performing model was Random Forest Regressor with 98.3% and 92.9% accuracy for both height and leaves. The least performance was observed with the Bayesian Ridge regressor model. Though the accuracy score is better than the existing system, the existing system input and output parameters is different compared to the current study. However, since the growth is predicted based on multiple parameters, the co-variance between the parameters is considered for prediction. The error rate can be reduced by considering a larger dataset and more features to better predict the yield.

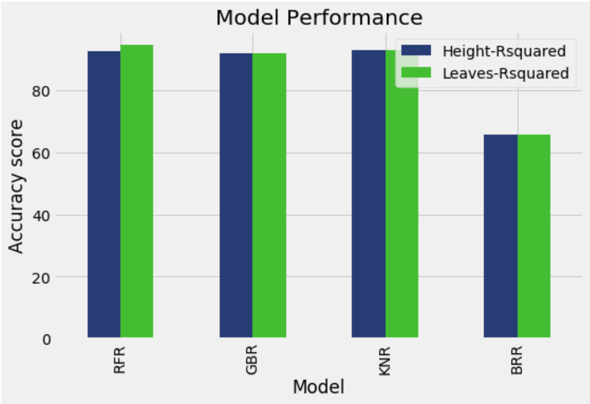


Fig. 7. Comparison of regression models- accuracy score.

The *K*- Nearest neighbour and Random Forest regressor models turn out to be the best model with the least errors. Figure 8 demonstrates the comparison of models with the least Root mean square errors for predicting the height and leaves of the plant. The prediction accuracy score of the present models was more remarkable compared to the existing models discussed in the literature. The leafy vegetables in hydroponic can grow faster when pH is maintained within an optimal range, and predictions of the growth can be improved if light intensity, dissolved oxygen of the solution, CO_2 can also be considered in the analysis.

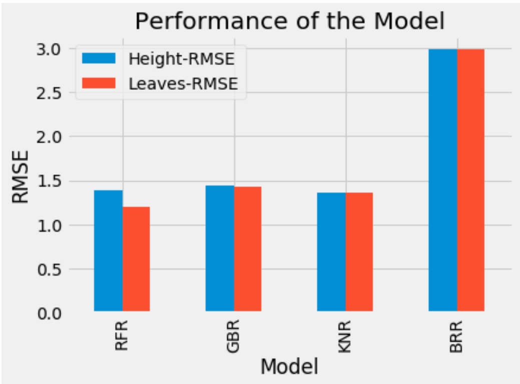


Fig. 8. Comparison of regression models- RMS errors.

4. Conclusion and Future Work

Numerous studies have been carried out to predict the future crops with a smart farming hydroponic system to implement machine learning models. The raft system for the spinach plant was cultivated under artificial grow lights. The goal of this study is to analyse the correlation between the input and output parameters and also to predict plant growth.

The Spearman correlation model echoes the degree of concordance and discordances between the parameters with rank-based observations. Spearman and Kendall's models can be implemented with a smaller dataset for both continuous and ordinal scales. These two are more robust and statistically efficient than any other correlation analysis models. Among the environment and nutrient solution parameters, the strongest correlation was found to be the solution's EC. The pH value has affected plant growth with a decreasing relationship. However, the other parameters had a lesser impact on plant growth in height and number of leaves.

Further, the regression models were compared with their prediction accuracy scores or co-efficient of determination. The accuracy of the Random Forest Regressor model became a good performance indicator. The Random Forest and *K*-Nearest Neighbor regressor models showed lower RMS errors compared to other regression models. This study's limitations were that the data collected samples could be tested further with other prediction models with a larger dataset. For the future prediction of crop yield with hydroponics, different plant species can evaluate prediction accuracy. Also, for precise predictions, the plants can be treated by varying EC and other treatments of nutrient solution temperature.

References

1. MarketandMarket. (2020). Hydroponics market by type, crop type, equipment, input, region – global forecast to 2025. Retrieved December 1, 2020, from https://www.marketsandmarkets.com/Market-Reports/hydroponic-market-94055021.html.
2. Solis-Toapanta, E.; Fisher, P.; and Gómez, C. (2020). Growth rate and nutrient uptake of basil in small-scale hydroponics. *HortScience*, 55(4), 507-514.
3. Caparrotta, S.; Masi, E.; Atzori, G.; Diamanti, I.; Azzarello, E.; Mancuso, S.; and Pandolfi, C. (2019). Growing spinach (*Spinacia oleracea*) with different seawater concentrations: effects on fresh, boiled and steamed leaves. *Scientia Horticulturae*, 256.
4. Gent, M.P. (2017). Factors affecting relative growth rate of lettuce and spinach in hydroponics in a greenhouse. *HortScience*, 52(12), 1742-1747.
5. Sago, Y.; and Shigemura, A. (2018). Quantitative nutrient management reduces nitrate accumulation in hydroponic butterhead lettuces grown under artificial lighting. HortScience, 53(7), 963-967.
6. Wortman, S.E. (2015). Crop physiological response to nutrient solution electrical conductivity and pH in an ebb-and-flow hydroponic system. Scientia Horticulturae, 194, 34-42.
7. Cho, Y.Y.; Oh, S.; Oh, M.M.; and Son, J.E. (2007). Estimation of individual leaf area, fresh weight, and dry weight of hydroponically grown cucumbers (Cucumis sativus L.) using leaf length, width, and SPAD value. Scientia *H*orticulturae, 111(4), 330-334.
8. Kocian, A.; Massa, D.; Cannazzaro, S.; Incrocci, L.; Di Lonardo, S.; Milazzo, P.; and Chessa, S. (2020). Dynamic bayesian network for crop growth prediction in greenhouses. Computers and Electronics in Agriculture, 169.
9. Lenni, L.; Suhardiyanto, H.; Seminar, K.B.; and Setiawan, R.P. (2020).

Photosynthetic rate of lettuce cultivated on floating raft hydroponic with controlled nutrient solution. HAYATI Journal of Biosciences, 27(1).

10. Qaddoum, K.; Hines, E.L.; and Iliescu, D.D. (2013). Yield prediction for tomato greenhouse using EFuNN. International Scholarly Research Notices. 2013.
11. Nagano, S.; Moriyuki, S.; Wakamori, K.; Mineno, H.; and Fukuda, H. (2019). Leaf-movement-based growth prediction model using optical flow analysis and machine learning in plant factory. Frontiers in *Plant Science*, 10.
12. Saputra, T.W.; Masithoh, R.E.; and Achmad, B. (2017). Development of plant growth monitoring system using image processing techniques based on multiple images. Proceeding of the 1st International Conference on Tropical Agriculture. Springer, 647-653.
13. Gillespie, D.P.; Kubota, C.; and Miller, S.A. (2020). Effects of low pH of hydroponic nutrient solution on plant growth, nutrient uptake, and root rot disease incidence of basil (*Ocimum basilicum L.*). HortScience, 55(8), 1251-1258.
14. Shadrin, D.; Somov, A.; Podladchikova, T.; and Gerzer, R. (2018). Pervasive agriculture: measuring and predicting plant growth using statistics and 2D/3D imaging. *Proceedings of the* International Instrumentation and Measurement Technology Conference. Houston, USA, 1-6.
15. Pabiania, M.D.; Caluyo, F.S.; and Linsangan, N.B. (2011). Wireless data acquisition and pH and conductivity levels prediction using genetic algorithm for hydroponics. *Proceedings of the* World Congress on Engineering and Computer Science. San Francisco, USA, 23-28.
16. Puth, M.T.; Neuhäuser, M.; and Ruxton, G.D. (2015). Effective use of Spearman's and Kendall's correlation coefficients for association between two measured traits. Animal Behaviour, 102, 77-84.
17. Segal, M.; and Xiao, Y. (2011). Multivariate random forests. W*IREs* Data *Mining and Knowledge Discovery*, 1(1), 80-87.
18. Zhang, Y.; and Haghani, A. (2015). A gradient boosting method to improve travel time prediction. Transportation Research Part C: Emerging Technologies, 58(B), 308-324.
19. Saqib, M. (2020). Forecasting COVID-19 outbreak progression using hybrid polynomial-Bayesian ridge regression model. Applied Intelligence, 51, 2703-2713.